Chapter 8

Microcanonical ensemble

8.1 Definition

We consider an *isolated* system with $\boxed{N}$ particles and energy $\boxed{E}$ in a volume $\boxed{V}$. By definition, such a system exchanges neither particles nor energy with the surroundings.

Uniform distribution of microstates. The assumption, that thermal equilibrium implies that the distribution function $\rho(q,p)$ of the system is a function of its energy,

$$\rho(q,p) \;=\; \rho(H(q,p)), \qquad\qquad \frac{d}{dt}\rho(q,p) \;=\; \frac{\partial\rho}{\partial H}\,\dot{E} \;\equiv\; 0 \;,$$

leads to to a constant $\rho(q,p)$, which is manifestly consistent with the ergodic hypothesis and the postulate of a priori equal probabilities discussed in Sect. 7.5.

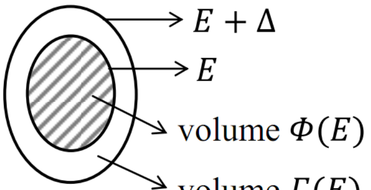
Energy shell. We consider a small but finite shell $[E, E+\Delta]$ close to the energy surface. The *microcanonical ensemble* is then defined by

$$\boxed{\rho(q,p) \;=\; \left\{ \begin{array}{ll} \dfrac{1}{\Gamma(E,V,N)} & \qquad E < H(q,p) < E+\Delta \\ 0 & \qquad \text{otherwise} \end{array} \right.} \tag{8.1}$$

microcanonical ensemble

We defined in (8.1) with

$$\Gamma(E,V,N) \;=\; \int\!\!\int_{E<H(q,p)<E+\Delta} d^{3N}q\; d^{3N}p \tag{8.2}$$



the volume occupied by the microcanonical ensemble. This is the volume of the shell bounded by the two energy surfaces with energies E and $E+\Delta$

The dependence on the spatial volume V comes (8.2) from the limits of the integration over dq_i.

Infinitesimal shell. Let $\Phi(E,V)$ be the total volume of phase space enclosed by the energy surface E. We then have that

$$\Gamma(E) \;=\; \Phi(E+\Delta) - \Phi(E) \;.$$

Taking with $\Delta \to 0$ the limit of an infinitesimal shell thickness we obtain

$$\Gamma(E) \;=\; \frac{\partial\Phi(E)}{\partial E}\Delta \;\equiv\; \Omega(E)\,\Delta, \qquad\qquad \Delta \ll E\,. \tag{8.3}$$

Density of states. The quantity $\Omega(E)$,

$$\boxed{\Omega(E) \;=\; \frac{\partial\Phi(E)}{\partial E} \;=\; \int d^{3N}q \int d^{3N}p\; \delta(E - H(q,p))}\,, \tag{8.4}$$

is the *density of states* at energy E. The distribution of microstates ρ is therefore given in terms of $\Omega(E)$, as

$$\rho(q,p) \;=\; \begin{cases} \dfrac{1}{\Omega(E)\Delta} & E < H(q,p) < E + \Delta \\ 0 & \text{otherwise} \end{cases}$$

Note that the formal divergence of $\rho(q,p)$ with $1/\Delta$. We will discuss in Sect. 8.2.4 that the actual value of Δ is not of importance.

8.2 Entropy

The expectation value of a classical observable $O(q,p)$ can be obtained by averaging over the probability density $\rho(q,p)$ of the microcanonical ensemble,

$$\langle O\rangle \;=\; \int d^{3N}q\, d^{3N}p\; \rho(q,p)\; O(q,p) \;=\; \frac{1}{\Gamma(E,V,N)} \int\!\!\int_{E<H(q,p)<E+\Delta} d^{3N}q\; d^{3N}p\; O(q,p)$$

The *entropy* can however not been be obtained as an average of a classical observable. It is instead a function of the overall number of available states.

POSTULATE The entropy is, according to Boltzmann, proportional to the logarithm of the number of available states, as measured by the phase space volume Γ:

$$S \;=\; k_B \ln\left(\frac{\Gamma(E,V,N)}{\Gamma_0(N)}\right) \tag{8.5}$$

We will now discuss the ramification of this definition.

Normalization factor. The normalization constant $\Gamma_0(N)$ introduced in (8.5) has two functionalities.

- $\Gamma_0(N)$ cancels the dimensions of $\Gamma(E,V,N)$. The argument of the logarithm is consequently dimensionless.

- The number of particles N is one of the fundamental thermodynamic variables. The functional dependence of $\Gamma_0(N)$ on N is hence important.

Incompleteness of classical statistics. It is not possible to determine $\Gamma_0(N)$ correctly within classical statistics. We will derive lateron that $\Gamma_0(N)$ takes the value

$$\boxed{\Gamma_0(N) \;=\; h^{3N} N!}\;, \tag{8.6}$$

in quantum statistics, where $h = 6.63 \cdot 10^{-34}\,\mathrm{m}^2\mathrm{kg/s}$ is the Planck constant. We consider this value also for classical statistics.

- The factor h^{3N} defines the reference measure in phase space which has the dimensions $[q]^{3N}[p]^{3N}$. Note that $[h] = \mathrm{m}\,(\mathrm{kg\,m/s}) = [q][p]$.
- $N!$ is the counting factor for states obtained by permuting particles.

> The factor $N!$ arises from the fact that particles are *indistinguishable.*
> Even though one may be in a temperature and density range where the motion of molecules can be treated to a very good approximation by classical mechanics, one cannot go so far as to disregard the essential indistinguishability of the molecules; one cannot observe and label individual atomic particles as though they were macroscopic billiard balls.

We will discuss in Sect. 8.2.1 the *Gibbs paradox*, which arises when on regards the constituent particles as distinguishable. In this case there would be no factor $N!$ in $\Gamma_0(N)$.

The entropy as an expectation value. We rewrite the definition (8.5) of the entropy as

$$S \;=\; -k_B \ln\left(\frac{\Gamma_0(N)}{\Gamma(E,V,N)}\right) \;=\; -k_B \int\!\!\int d^{3N}q\; d^{3N}p\; \rho(q,p) \ln\big[\Gamma_0(N)\rho(q,p)\big]\;, \tag{8.7}$$

where we have used that $\rho(q,p) = 1/\Gamma(E,V,N)$ within the energy shell and that

$$\int\!\!\int d^{3N}q\; d^{3N}p\; \rho(q,p) \;=\; \int_{E<H<E+\Delta} d^{3N}q\; d^{3N}p/\Gamma(E,V,N) \;=\; 1\;.$$

We hence have

$$\boxed{S \;=\; -k_B \big\langle \ln[\Gamma_0(N)\rho(q,p)]\big\rangle}\;.$$

The entropy coincides hence with Shannon's information-theoretical definition of the entropy, apart from the factors k_B and $\Gamma_0(N)$, as discussed in Sect. 5.5.3.

Thermodynamic consistency. Since we have introduced the entropy definition in an ad-hoc way, we need to convince ourselves that it describes the thermodynamic entropy as a state function. The entropy must therefore fulfill the requirements of

(1) additivity;

(2) consistency with the definition of the temperature;

(3) consistency with the second law of thermodynamics;

(4) adiabatic invariance.

8.2.1 Additivity; Gibbs paradox

The classical Hamiltonian $H(q,p) = H_{kin}(p) + H_{int}(q)$ is the sum of the kinetic energy $H_{kin}(q)$ and of the particle-particle interaction $H_{int}(q)$. The condition

$$E < H(q,p) < E + \Delta \qquad\qquad E < H_{kin}(p) + H_{int}(q) < E + \Delta$$

limiting the available phase space volume $\Gamma(E,V,N)$ on the energy shell, as defined by (8.2) could then be fulfilled by a range of combinations of $H_{kin}(p)$ and $H_{int}(q)$.

The law of large numbers, which we will discuss in Sect. 8.6, implies however that both the kinetic and the interaction energies take well defined values for large particle numbers N.

Scaling of the available phase space. The interaction between particles involves only pairs of particles, with the remaining $N-2 \approx N$ particles moving freely within the available volume V. This consideration suggest together with an equivalent argument for the kinetic energy that the volume of the energy shell scales like

$$\int\int_{E<H(q,p)<E+\Delta} d^{3N}q\, d^{3N}p \;=\; \boxed{\Gamma(E,V,N) \;\sim\; V^N\,\gamma^N(E/N,V/N)}\,. \tag{8.8}$$

We will verify this relation in Sect. 8.4 for the classical ideal gas.

Extensive entropy. Using scaling relation (8.8) for the volume of the energy shell and the assumption that $\Gamma_0(N) = h^{3N}N!$ we then find that the entropy defined by (8.5) is extensive:

$$\begin{aligned} S &= k_B \ln\left(\frac{\Gamma(E,V,N)}{\Gamma_0(N)}\right) = k_B \ln\left(\frac{V^N\gamma^N(E/N,V/N)}{h^{3N}N!}\right) \\ &= k_B\, N\left[\ln\left(\frac{V}{N}\frac{\gamma}{h^3}\right)+1\right] \equiv k_B\, N\, s(E/N,V/N) \end{aligned} \tag{8.9}$$

where we have used the Stirling formula $N! \approx \sqrt{2\pi N}(N/e)^N$, namely that $\ln(N!) \approx N(\ln(N)-1)$.

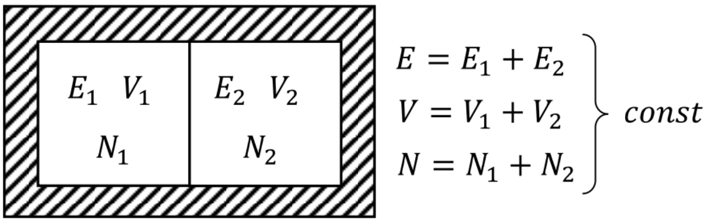
Gibbs paradox. The extensivity of the entropy result in (8.9) from the fact that $V^N/N! \approx (V/N)^N$. Without the factor $N!$ in $\Gamma_0(N)$, which is however not justifiable within classical statistics, the entropy would not be extensive. This is the *Gibbs paradox.*

Additivity in the case of identical thermodynamic states. Two systems have identical thermodynamic properties if they intensive variables are the same, viz temperature

T, pressure P, particle density N/V and energy density E/N. It then follows directly from (8.9) that

$$S(E,V,N) \;=\; k_B\,(N_1+N_2)\, s(E/N,V/N) \;=\; S(E_1,V_1,N_1)+S(E_2,V_2,N_2)\ ,$$

where we have used that $N=N_1+N_2$.



Additivity for two systems in thermal contact. Two systems defined by E_1, V_1, N_1 and respectively E_2, V_2, N_2 in thermal contact may allow energy such that the total energy $E=E_1+E_2$ is constant. For the argument of the entropy we have then:

$$\frac{\Gamma(E,V,N)}{\Gamma_0(N)} \;=\; \sum_{E_1} \frac{\Gamma(E_1,V_1,N_1)}{\Gamma_0(N_1)}\,\frac{\Gamma(E-E_1,V_2,N_2)}{\Gamma_0(N_2)} \tag{8.10}$$

The law of large numbers tells us that the right-hand-side is sharply peaked at its maximum value $E_1=E_{max}$ and that the width of the peak has a width scaling with $\sqrt{E_{max}}$. We hence have

$$S(E_{max}) \;<\; S(E,N,V) \;<\; k_B \ln\left(\sqrt{E_{max}}\right)+S(E_{max})\ , \tag{8.11}$$

where the first inequality is due to the fact that a single term is smaller than the sum of positive terms. The second inequality in (8.11) results when when one replaces the sum on the r.h.s. of (8.10) by the product of the width $\sqrt{E_{max}}$ of the peak and its height. We have defined in (8.11)

$$S(E_{max}) \;=\; k_B \ln\left(\frac{\Gamma(E_{max},V_1,N_1)\,\Gamma(E-E_{max},V_2,N_2)}{\Gamma_0(N_1)\,\Gamma_0(N_2)}\right)\ ,$$

from which follows that $S(E,V,N) \;=\; S(E_1,V_1,N_1)+S(E_2,V_2,N_2)$. Note that the entropy $S(E_{max})$ is extensive and that the term $\sim \ln(E_{max})$ in (8.11) is hence negligible in the thermodynamic limit $N\to\infty$.

8.2.2 Consistency with the definition of the temperature

Two systems with entropies $S_1=S(E_1,V_1,N_1)$ and $S_2=S(E_2,V_2,N_2)$ in thermal contact may exchange energy in the form of heat, with the total entropy,

$$0 \;=\; dS \;=\; \frac{\partial S_1}{\partial E_1}dE_1+\frac{\partial S_2}{\partial E_2}dE_1, \qquad dE_1 \;=\; -dE_2\ , \tag{8.12}$$

becoming stationary at equilibrium. Note that the total energy $E_1 + E_2$ is constant. The equilibrium condition (8.12) implies that there exists a quantity T, denoted *temperature*, such that

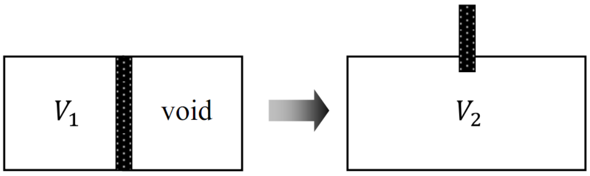
$$\boxed{\frac{\partial S_1}{\partial E_1} = \frac{1}{T} = \frac{\partial S_2}{\partial E_2}} \,. \qquad (8.13)$$

The possibility to define the temperature, as above, is hence a direct consequence of the conservation of the total energy. From the microcanonical definition of the entropy one only needs that the entropy is a function only of the internal energy, via the volume $\Gamma(E, V, N)$ of the energy shell, and not of the underlying microscopic equation of motion.

8.2.3 Consistency with the second law of thermodynamics

We would like to confirm here that the statistical entropy, defined by (8.5) fulfills the *second law of thermodynamics*

> *"If an isolated system undergoes a process between two states at equilibrium, the entropy of the final state cannot be smaller than that of the initial state."*



Free expansion. Both the energy E and and the number of particle stays constant during a free expansion, defined by the absence of external heat transfer,

$$\Delta Q \;=\; 0 \,.$$

The volume $\Gamma(E, V, N)$ of the energy shell increase with increasing volume V. We hence have

$$\boxed{S(E, V_2, N) \;>\; S(E, V_1, N)}\,, \qquad S \;=\; k_B \ln\left(\frac{\Gamma(E,V,N)}{\Gamma_0(N)}\right), \qquad V_2 \;>\; V_1 \,,$$

where have made use of the fact that the normalization factor $\Gamma_0(N)$ remains constant.

Dynamical constraints. Dynamical constraints (viz bouncing from the wall) are mitigated when the volume is increased.

EQUIVALENCE	The second law is equivalent to saying that the entropy rises when dynamical constraints are eliminated,

8.2.4 How thick is the energy shell?

The definition Sect. 8.5 of the entropy involves the volume $\Gamma(E, V, N)$ of a shell of state space of width Δ centered around the energy E. It seems therefore that the entropy $S = S_\Delta(E, V, N)$ depends on an unspecified parameter Δ. Is the entropy then not uniquely specified?

Reference energy. For small Δ we may use the approximation

$$\Gamma(E,V,N) \approx \Omega(E)\,\Delta, \qquad \Omega(E) = \frac{\partial\Phi(E)}{\partial E}\ , \tag{8.14}$$

where $\Omega(E)$ is the density of states, as defined previously in (8.3). In order to decide whether a given Δ is small or large we need a reference energy Δ_0. One may take f.i. $\Delta_0 \sim k_B T$, which corresponds, as shown in Sect. 3.4.1, in order of magnitude to the thermal energy of an individual particle.

Thermodynamic limit. The entropy involves the logarithm of $\Gamma(E,V,N)$,

$$\ln\left(\frac{\Gamma(E,V,N)}{\Gamma_0}\right) = \ln\left(\frac{\Omega(E)\Delta\Delta_0}{\Gamma_0\Delta_0}\right) = \underbrace{\ln\left(\frac{\Omega(E)\Delta_0}{\Gamma_0}\right)}_{\propto N} + \ln(\Delta/\Delta_0)\ ,$$

where we have taken care the arguments of the logarithms are dimensionless. The key insight resulting from this representation is that the exact value of both Δ and Δ_0 is irrelevant in the thermodynamic limit $N\to\infty$ as long as

$$\big|\ln(\Delta/\Delta_0)\big| \ll N\ .$$

Energy quantisation will ensure this condition in quantum statistics.

Shell vs. sphere. We may also considering the limit of large Δ to the extend that we may substitute the volume $\Phi(E)$ for the the volume $\Gamma(E,V,N)$ of the energy shell,

$$\ln\big(\Gamma(E,V,N)\big) \approx \ln\big(\Phi(E)\big)\ , \tag{8.15}$$

where we have disregarded the normalization factor Γ_0. The reason that (8.15) holds stems from the fact that the volume and surface of a sphere with radius R of dimension $3N$, compare Sect. 8.3.1, scale respectively like R^{3N} and R^{3N-1}. This scaling leads to

$$\begin{aligned} \ln\big(\Phi(E)\big) &\sim \ln\big(R^{3N}\big) = 3N\ln(R) \\ \ln\big(\Gamma(E,V,N)\big) &\sim \ln\big(R^{3N-1}\Delta\big) = \underbrace{(3N-1)\ln(R) + \ln(\Delta)}_{\sim 3N\ln(R)}\ , \end{aligned}$$

which corroborates (8.15). We did not perform here an analysis of the units involved, neglected in particular the reference energy Δ_0.

8.3 Calculating with the microcanonical ensemble

In order to perform calculations in statistical physics one proceeds through the following steps.

1) Formulation of the Hamilton function

$$H(q,p) = H(q_1,\dots,q_{3N},p_1,\dots,p_{3N},z)\ ,$$

where z is some external parameter, e.g., volume V. $H(q,p)$ specifies the microscopic interactions.

2) Determination of the phase space $\Gamma(E,V,N)$ and calculation of the density of states $\Omega(E,V,N)$:

$$\Omega(E,V,N) \;=\; \int d^{3N}q \int d^{3N}p \; \delta(E - H(q,p)) \; .$$

3) Calculation of the entropy from the volume $\Phi(E)$ of the energy sphere:

$$S(E,V,N) \;=\; k_B \ln\left(\frac{\Phi(E)}{\Gamma_0}\right) \; .$$

4) Calculation of P, T, μ:

$$\frac{1}{T} \;=\; \left(\frac{\partial S}{\partial E}\right)_{V,N} , \qquad -\frac{\mu}{T} \;=\; \left(\frac{\partial S}{\partial N}\right)_{E,V} , \qquad \frac{P}{T} \;=\; \left(\frac{\partial S}{\partial V}\right)_{E,N} \; .$$

5) Calculation of the internal energy:

$$U \;=\; \langle H\rangle \;=\; E(S,V,N) \; .$$

6) Calculation of other thermodynamic potentials and their derivatives by application of the Legendre transformation:

$$\begin{aligned} F(T,V,N) &= U - TS, \\ H(S,P,N) &= U + PV, \\ G(T,P,N) &= U + PV - TS \; . \end{aligned}$$

7) One can calculate other quantities than the thermodynamic potentials, for instance, probability distribution functions of certain properties of the system, e.g., momenta/velocity distribution functions. If the phase space density of a system of N particles is given by

$$\rho(q,p) \;=\; \rho(\vec{q}_1,\ldots,\vec{q}_N,\vec{p}_1,\ldots,\vec{p}_N) \; ,$$

then the probability of finding particle i with momentum $\vec{p}$ is

$$\begin{aligned} \rho_i(\vec{p}) &= \langle \delta\left(\vec{p} - \vec{p}_i\right)\rangle \\ &= \int d^3q_1 \ldots d^3q_N \int d^3p_1 \ldots d^3p_N \; \rho\left(\vec{q}_1,\ldots,\vec{q}_N,\vec{p}_1,..,\vec{p}_i,..,\vec{p}_N\right) \delta\left(\vec{p} - \vec{p}_i\right) \; . \end{aligned}$$

8.3.1 Hyperspheres

Let us calculate for later purposes the volume

$$\Omega_n(R) \;=\; \int_{\sum_{i=1}^{n} x_i^2 < R^2} d^n x \;=\; R^n \Omega_n(1) \tag{8.16}$$

of a hypersphere of n dimensions and radius P.

Spherical coordinates. We notice that the volume $\Omega_n(1)$ of the sphere with unity radius enters the determinant of the Jacobian when transforming via

$$d^n x \,=\, dx_1 \dots dx_n \,=\, \Omega_n(1)\, nR^{n-1}\, dR \tag{8.17}$$

euclidean to spherical coordinates. This transformation is valid if the integrand depends exclusively on the radius R.

Gaussian integrals. In order to evaluate (8.16) we make use of the fact that we can rewrite the Gaussian integral

$$\int_{-\infty}^{+\infty} dx\, e^{-x^2} \,=\, \sqrt{\pi}, \qquad\qquad \int_{-\infty}^{+\infty} dx_1 \dots \int_{-\infty}^{+\infty} dx_n\, e^{-(x_1^2+\ldots+x_N^2)} \,=\, \pi^{n/2}\ .$$

as

$$\begin{aligned}\pi^{n/2} \;&=\; \int_0^\infty e^{-R^2}\, \Omega_n(1)\, nR^{n-1}\, dR \,=\, n\Omega_n(1) \int_0^\infty e^{-y}\, y^{(n-1)/2}\, \frac{dy}{2\sqrt{y}} \\ &=\; \frac{n}{2}\Omega_n(1) \int_0^\infty e^{-y}\, y^{\frac{n}{2}-1}\, dy\ ,\end{aligned} \tag{8.18}$$

where we have used (8.17), $\sum_i x_i^2 = R^2 \equiv y$ and that $2RdR = dy$.

Gamma function. With the definition

$$\Gamma(z) \,=\, \int_0^\infty dx\, x^{z-1} e^{-x},$$

of the Γ-function we then obtain from (8.18) that

$$\pi^{n/2} \,=\, \frac{n}{2}\Omega_n(1)\Gamma(n/2), \qquad\qquad \Omega_n(1) \,=\, \frac{\pi^{n/2}}{(n/2)\,\Gamma(n/2)}\ . \tag{8.19}$$

Note that we did evaluate the volume of a hypersphere for formally dimensionless variables x_i. For later purposes we will need that

$$\Gamma(N) \,=\, (N-1)\Gamma(N-1), \qquad\qquad \Gamma(N) \,=\, (N-1)! \tag{8.20}$$

whenever the argument of the Γ-function is an integer.

8.4 The classical ideal gas

We consider now the steps given in Sect. 8.3 in order to analyze an ideal gas of N particles in a volume V, defined by the Hamilton function

$$H(q,p) \,=\, \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m}\ ,$$

where m is the mass of the particles.

Phase space volume. We will make use of (8.15), namely that the volume $\Gamma(E,V,N)$ of the energy sphere

$$E \;<\; H \;<\; E+\Delta.$$

can be replaced by the volume of the energy shell,

$$\Phi(E) \;=\; \int\!\!\int_{\sum_{i=1}^{3N} p_i^2\le 2mE} d^{3N}q\, d^{3N}p \;=\; V^N \underbrace{\int_{\sum_{i=1}^{3N} p_i^2\le 2mE} d^{3N}p}_{\Omega_{3N}(\sqrt{2mE})} \;, \tag{8.21}$$

when it comes to calculating the entropy in the thermodynamic limit. We have identified the last integral in (8.21) as the volume of a $3N$-dimensional sphere with radius $\sqrt{2mE}$.

Using (8.16) and (8.19),

$$\Omega_{3N}(\sqrt{2mE}) \;=\; (2mE)^{3N/2}\,\Omega_{3N}(1), \qquad \Omega_{3N}(1) \;=\; \frac{\pi^{3N/2}}{(3N/2)\,\Gamma(3N/2)} \;,$$

we obtain

$$\boxed{\Phi(E) \;=\; V^N \Omega_{3N}(1)\left(\sqrt{2mE}\right)^{3N}} \;. \tag{8.22}$$

8.4.1 Entropy

Using (8.22) we find

$$S(E,V,N) \;=\; k_B \ln\left(\frac{\Phi(E)}{h^{3N}N!}\right) \;=\; \boxed{k_B \ln\left(\frac{V^N\left(\sqrt{2mE}\right)^{3N}\;\Omega_{3N}(1)}{h^{3N}N!}\right)} \tag{8.23}$$

for the entropy of a classical gas. It is easy to check, that the argument of the logarithm is dimensionless as it should be.

Large N expansion. For $N >> 1$, one may use the Stirling formula, $N! \approx \sqrt{2\pi N}(N/e)^N$, to expand the Γ-function for integer argument as

$$\begin{aligned} \ln\big(\Gamma(N)\big) \;&=\; \ln\big((N-1)!\big) \\ &\approx\; (N-1)\ln(N-1) - (N-1) \\ &\approx\; N\ln(N) - N \;, \end{aligned}$$

in order to simplify the expression for $S(E,V,N)$. Using (8.19) we perform the following algebraic transformations to $\Omega_{3N}(1)$:

$$\begin{aligned}\ln\Omega_{3N}(1) &= \ln\left(\frac{\pi^{\frac{3N}{2}}}{\frac{3N}{2}\Gamma\left(\frac{3N}{2}\right)}\right)\\ &= \frac{3N}{2}\ln\pi-\left[\frac{3N}{2}\ln\left(\frac{3N}{2}\right)-\frac{3N}{2}\right]\\ &= \frac{3N}{2}\ln\frac{2\pi}{3N}+\frac{3N}{2}\\ &= N\left[\ln\left(\frac{2\pi}{3N}\right)^{3/2}+\frac{3}{2}+\mathcal{O}\left(\frac{\ln N}{N}\right)\right].\end{aligned}$$

We insert this expression in Eq. (8.23) and obtain:

$$S = k_B N\left\{\ln\left(\frac{V(2mE)^{3/2}}{h^3}\right)+\ln\left(\frac{2\pi}{3N}\right)^{3/2}+\frac{3}{2}-\underbrace{(\ln N-1)}_{\frac{\ln N!}{N}}\right\}. \tag{8.24}$$

Sackur-Tetrode Equation. Rewriting (8.24) as

$$\boxed{S = k_B N\left\{\ln\left[\left(\frac{4\pi mE}{3h^2N}\right)^{3/2}\frac{V}{N}\right]+\frac{5}{2}\right\}} \tag{8.25}$$

we obtain the *Sackur-Tetrode equation.*

Equations of state. Now we can differentiate the Sackur-Tetrode equation (8.25) and thus obtain

- the *caloric equation* of state (3.5) for the ideal gas:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N} = Nk_B\frac{3}{2}\frac{1}{E}, \qquad \boxed{E = \frac{3}{2}Nk_BT = U}$$

- the *thermal equation* of state for the ideal gas:

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V}\right)_{E,T} = \frac{k_BN}{V}, \qquad \boxed{PV = Nk_BT}.$$

Gibbs Paradox. If we hadn't considered the factor $N!$ when working out the entropy, then one would obtain that

$$S_{classical} = k_B N\left\{\ln\left[\left(\frac{4\pi mE}{3h^2N}\right)^{3/2}V\right]+\frac{3}{2}\right\}.$$

With this definition, the entropy is *non-additive*, i.e.,

$$S(E,V,N) \;=\; Ns\left(\frac{E}{N},\frac{V}{N}\right)$$

is not fulfilled, as mentioned already in Sect. 8.2.1. This was realized by *Gibbs*, who introduced the factor $N!$ and attributed it to the fact that the particles are indistinguishable.

8.5 Fluctuations and correlation functions

We use, as defined in Sect. 7.2.3, the probability distribution function $\rho_M(q,p)$ in phase space, which could be evaluated as a time average

$$\rho(q,p) \;=\; \lim_{M\to\infty}\frac{1}{M}\sum_{l=1}^{M}\delta^{(3N)}\left(q-q^{(l)}\right)\delta^{(3N)}\left(p-p^{(l)}\right)$$

of M measurements of the states $(q^{(l)},p^{(l)})$ visited along a given trajectory. We have pointed out in (8.1) that $\rho(q,p)$ is uniformly distributed on the energy shell for a microcanonical ensemble.

Observables. A classical observable $B(q,p)$ corresponds to a function

$$B(q,p) \;=\; B\left(q_1,\ldots,q_{3N},p_1,\ldots,p_{3N}\right),$$

phase state. We start with the basic properties.

Expectation value. The *average value* of $B(q,p)$ is

$$\begin{aligned}\langle B(q,p)\rangle \;&=\; \lim_{M\to\infty}\frac{1}{M}\sum_{l=1}^{M}B\left(q^{(l)},p^{(l)}\right)\\ &=\; \int_\Gamma d^{3N}q\; d^{3N}p\;\rho(q,p)\;B(q,p)\;.\end{aligned}$$

Variance. The *Variance* of a function on phase space is

$$(\Delta B)^2 \;=\; \left\langle (B-\langle B\rangle)^2\right\rangle \;=\; \left\langle B^2\right\rangle-\langle B\rangle^2 \;=\; \sigma_B^2\;.$$

The variance provides information about the fluctuation effects around the average value.

Higher momenta. A *momentum* of order n is defined generically as

$$\mu_n \;=\; \langle B^n\rangle\;.$$

Probability distribution of an observable. We may define with

$$\rho(b) \;=\; \int d^{3N}q\; d^{3N}p\;\rho(q,p)\;\delta(b-B(q,p)) \;=\; \langle\delta(b-B(q,p))\rangle \tag{8.26}$$

also the probability distribution $\rho(b)$ of the observable B. The expectation value of B is then given as

$$\langle B\rangle \;=\; \int db\; b\;\rho(b)\;.$$

8.5.1 Characteristic function and cumulants

The characteristic function of an observable

$$\boxed{\varphi_B(z) \;=\; \left\langle e^{izB}\right\rangle} \;=\; \int d^{3N}q\, d^{3N}p\, \rho(q,p)\, e^{izB(q,p)}\ . \tag{8.27}$$

is both a function of the dummy variable[*] z and a *functional*[†] of the observable $B = B(q,p)$.

Generating functional. The characteristic function of a random variable B serves as a *generating functional*, as it allows via suitable differentions (with respect to the dummy variable z),

$$\mu_n \;=\; \frac{1}{i^n}\frac{d^n}{dz^n}\varphi_B(z)\Big|_{z=0}, \qquad \boxed{\varphi_B(z) \;=\; \sum_{n=0}^{\infty}\frac{(iz)^n}{n!}\mu_n}\ , \tag{8.28}$$

to recover the moments μ_n of underlying probability distribution.

Cummulants. In (8.27) we did expand $\varphi_B(z)$ in powers of z. Alternatively one may expand the logarithm of $\varphi_B(z)$ in powers of z. One obtains (here without proof)

$$\ln\varphi_B(z) \;=\; \sum_{n=1}^{\infty}\zeta_n\frac{(iz)^n}{n!} \;=\; \mu_1 iz - \sigma^2\frac{z^2}{2} + \ldots\ . \tag{8.29}$$

The coefficients

$$\begin{aligned}
\zeta_1 &= \mu_1 \;=\; \langle B\rangle \\
\zeta_2 &= \sigma^2 \;=\; \langle B^2\rangle - \langle B\rangle^2 \\
\zeta_3 &= \ldots
\end{aligned}$$

are called *cumulants.*

Extensivity. Cummulant expansions like (8.29) are found in many settings, they form, e.g., the basis of high-temperature expansions. The reason is that cumulants are (in contrast to the moments) consistent with extensivity requirements, f.i. with respect to the volume V:

$$\langle B\rangle \;\sim\; V, \qquad \langle B^2\rangle \;\sim\; V^2, \qquad \boxed{\zeta_2 \;=\; \langle B^2\rangle - \langle B\rangle^2 \;\sim\; N}\ .$$

All cumulants are size-extensive.

8.5.2 Correlations between observables

We define in analogy to (8.26) the *joint distribution function* $\rho(a,b)$,

$$\begin{aligned}
\rho(a,b) \;&=\; \big\langle\delta(a-A)\delta(b-B)\big\rangle \qquad\qquad (8.30)\\
&=\; \int d^{3N}q\, d^{3N}p\, \rho(q,p)\, \delta\big(a-A(q,p)\big)\, \delta\big(b-B(q,p)\big)
\end{aligned}$$

---

[*] A *dummy variable* doesn't have a physical meaning. It solely serves for computational purposes.

[†] A *functional* is a function of another function.

of two observables $A = A(q,p)$ and $B = B(q,p)$.

Cross correlation function. The joint distribution function $\rho(a,b)$ allows to investigate the cross correlations

$$\begin{aligned} C_{AB} &= \big\langle (A-<A>)(B-<B>) \big\rangle \\ &= \int da \int db\, \rho(a,b)\, \big(a - \langle A\rangle\big)\big(b - \langle B\rangle\big) \end{aligned} \tag{8.31}$$

between A and B.

Uncorrelated variables. Two observable are uncorrelated if they are *statistically independent.* The joint distribution factorizes is this case,

$$\rho(a,b) \;=\; \rho_A(a)\rho_B(b)\ .$$

The cross correlations vanish consequently:

$$\begin{aligned} C_{AB} &= \big\langle \big(A - \langle A\rangle\big)\big(B - \langle B\rangle\big)\big\rangle \\ &= \big\langle A - \langle A\rangle\big\rangle \big\langle B - \langle B\rangle\big\rangle \\ &= 0\ . \end{aligned}$$

8.6 Central limit theorem

Consider M statistically independent observables $B_m(q,p)$. The central limit theorem states then that the probability density of the observable

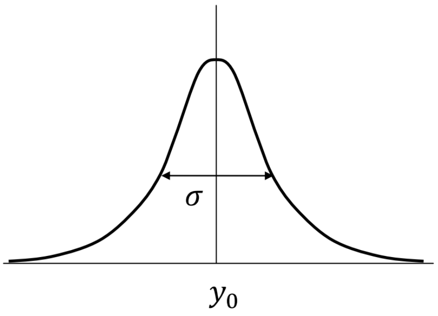
$$B(q,p) \;=\; \sum_{m=1}^{M} B_m(q,p)$$

is normal distributed with an extensive mean $\langle B\rangle \sim M$ and with the width proportional to $\sqrt{M}$.

8.6.1 Normal distribution

An observable Y is said to be normal distributed if the probability distribution function $\rho(y) = \langle \delta(y-Y)\rangle$ has the form of a Gauss curve,

$$\boxed{\rho(y) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-y_0)^2/(2\sigma^2)}} \tag{8.32}$$



The mean and the variance are

$$\langle Y\rangle \;=\; \int_{-\infty}^{+\infty} dy\, \rho(y) y \;=\; y_0\ ,$$

and

$$\big\langle (Y-<Y>)^2 \big\rangle \;=\; \int_{-\infty}^{+\infty} dy\, \rho(y)\, (y-y_0)^2 \;=\; \sigma^2$$

respectively. The even moments of the normal distribution are finite,

$$\left\langle Y^{2n}\right\rangle = \frac{(2n)!}{2^n n!}\sigma^{2n} ,$$

with the odd moments $\langle Y^{2n+1}\rangle = 0$ vanishing due to the reflection symmetry with respect to the mean y_0.

Characteristic function. The characteristic function of a normal distribution is also a Gauss function:

$$f(z) = \left\langle e^{izY}\right\rangle = \int_{-\infty}^{+\infty} dy\, e^{izy}\rho(y) = e^{izy_0}e^{-\frac{1}{2}z^2\sigma^2} . \tag{8.33}$$

For the proof consider the overall exponent

$$\begin{aligned} izy - \frac{(y-y_0)^2}{2\sigma^2} &= -\frac{y^2 - 2y_0y - 2izy\sigma^2 + y_0^2}{2\sigma^2} \\ &= -\frac{(y-y_0-iz\sigma^2)^2}{2\sigma^2} + izy_0 - \frac{1}{2}z^2\sigma^2 . \end{aligned}$$

8.6.2 Derivation of the central limit theorem

For the derivation of the central limit we define a variable Y as

$$Y \equiv \frac{1}{\sqrt{M}}\big(X_1 + X_2 + \ldots + X_M\big) ,$$

where X_1, X_2, …, X_M are M mutually independent variables such that $\langle X_i\rangle = 0$ for convenience.

Characteristic function of independent variables. We make now use of the fact, that the characteristic function of the sum independent variables, such as for Y, factorizes:

$$\begin{aligned} \left\langle e^{-izY}\right\rangle &= \left\langle e^{-iz(X_1+X_2+\ldots+X_M)/\sqrt{M}}\right\rangle \\ &= \left\langle \prod_{j=1}^{M} e^{-izX_j/\sqrt{M}}\right\rangle \\ &= \prod_{j=1}^{M}\left\langle e^{-izX_j/\sqrt{M}}\right\rangle . \end{aligned}$$

Using the representation

$$A_j(z/\sqrt{M}) \equiv \ln\left\langle e^{-izX_j/\sqrt{M}}\right\rangle \tag{8.34}$$

for the logarithm of the characteristic function $\langle \exp(-izX_j/\sqrt{M})\rangle$ of the individual random variables we then find

$$\boxed{\left\langle e^{-izY}\right\rangle = \exp\left(\sum_{j=1}^{M} A_j(z/\sqrt{M})\right)} . \tag{8.35}$$

Large M expansion. If M is a large number, $A_j(z/\sqrt{M})$ can be expanded as

$$\boxed{A_j(z/\sqrt{M}) \;\simeq\; A_j(0) + \frac{z^2}{2M}A_j''(0) + \frac{z^3}{M}\mathcal{O}(M^{-1/2})}\;. \tag{8.36}$$

Here, the first-order term $A_j'(0)$ in z vanishes because it is proportional to $\langle X_j\rangle = 0$.

Variance. The quadratic term is (8.36) is obtained by differentiating (8.34):

$$A_j''(0) \;=\; -\left\langle X_j^2\right\rangle \;.$$

Substituting now (8.36) into (8.34) one obtains with

$$\left\langle e^{-izY}\right\rangle \;=\; \exp\left\{-\frac{1}{2}z^2\sigma^2 + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)\right\}, \qquad \sigma^2 \;\equiv\; \frac{1}{M}\sum_{j=1}^{M}\left\langle X_j^2\right\rangle \tag{8.37}$$

the proof of the central limit theorem. Compare (8.37) with (8.33).

Finite mean. The expectation values add,

$$\langle Y\rangle \;=\; \sum_{j=1}^{M} < X_j > /\sqrt{M}\;,$$

when the means $\langle X_j\rangle \neq 0$ of the individual variables are finite. $Y - \langle Y\rangle$ is then normal distributed.

Binomial distribution. As a first example we consider the *binomial distribution* in the limit of a large number of draws (flipping of a coin):

$$n \;=\; \sum_{i=1}^{N} n_i \qquad\qquad n_i \;=\; \left\{\begin{array}{ll} 1 & \text{with probability} \quad p \\ 0 & \text{with probability} \quad 1-p \end{array}\right.$$

The probability $\rho_N(n)$ to get a one n times out of N coin flip is then

$$\boxed{\rho_N(n) \;=\; \left(\begin{array}{c} N \\ n \end{array}\right) p^n(1-p)^{N-n}}\;. \tag{8.38}$$

The central limit theorem predicts that

$$\boxed{\lim_{N\to\infty}\rho_N(n) \;\to\; \frac{1}{\sqrt{2\pi N\sigma^2}}e^{-(n-<n>)^2/2N\sigma^2}}\;, \tag{8.39}$$

where

$$\sigma^2 \;=\; \frac{1}{N}\sum_{i=1}^{N}\left(\langle n_i^2\rangle - \langle n_i\rangle^2\right) \;=\; p(1-p)$$

is the variance of a single coin flip.

Flat distribution. As a numerical example we consider the flat distribution on the interval $[0, 1]$. Already the sum of $N = 3$ draws is remarkably close to the large-N predictions of the central limit theorem.